Advanced Methods for Characterizing the Immersion Factor of Irradiance Sensors

STANFORD B. HOOKER

Laboratory for Hydrospheric Processes, NASA Goddard Space Flight Center, Greenbelt, Maryland

GIUSEPPE ZIBORDI

Institute for Environment Sustainability, Joint Research Centre, Ispra, Italy

(Manuscript received 3 October 2003, in final form 17 July 2004)

ABSTRACT

Two new immersion factor methods are evaluated by comparing them with the so-called traditional (or incremental) method. For the first method, the optical measurements taken at discrete water depths are substituted by continuous profiles created by removing the water from the tank used in the experimental procedure at a constant flow rate with a pump. In the second method, the commonly used large tank is replaced by a small water vessel with sidewall baffles, which permits the use of a quality-assured volume of water. The summary of the validation results produced for the different methods shows a significant convergence of the two new methods with the traditional method with differences generally well below 1%. The average repeatabilities for single-sensor characterizations (across seven wavelengths) of the three methods are very similar and approximately 0.5%. The evaluation of the continuous method demonstrates its full applicability in the determination of immersion factors with a significant time savings. The results obtained with the small water vessel demonstrate the possibility of significantly reducing the size of the tank (along with decreasing the execution time) and permitting a completely reproducible methodology (based on the use of pure water). The small tank approach readily permits the isolation and quantification of individual sources of uncertainty, the results of which confirm the following aspects of the general experimental methodology: (a) pure water is preferred over tap water, (b) the water should not be recycled (so it does not age), (c) bubbles should be removed from all wetted surfaces, (d) the water surface should be kept as clean as possible, (e) sidewall reflections can be properly minimized with internal baffles, and (f) a pure water characterization can be easily corrected to produce an appropriate seawater characterization. Within the context of experimental efficiency and reproducibility, this study suggests that the combination of a properly baffled small tank with a constant-flow pump would be an optimal system.

1. Introduction

The immersion factor $I_f(\lambda)$ is a necessary part of the spectral characterization of an in-water irradiance sensor (λ denotes wavelength), because when a cosine collector is immersed in water, its light transmissivity is less than it is in air. Irradiance sensors are calibrated in air, however, so a correction for this change must be applied when the in-water raw data are converted to physical units. The immersion factor must be determined experimentally, using a laboratory protocol, for each collector. When in situ measurements are used to create ground-truth databases for remote sensing calibration and validation activities, like those established for the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) project (Hooker and Esaias 1993), the uncertainties in the former ultimately influence the quality of the data products. The SeaWiFS ground-truth uncertainty budget can only be satisfied if each contributing uncertainty is on the order of 1%–2% (Hooker and McClain 2000). As a generalized description, this constitutes so-called 1% radiometry; in other words, uncertainty sources in the calibrated use of a sensor—like the immersion factor—must be kept at approximately the 1% level.

Studies of immersion effects date back to the work of Atkins and Poole (1933), who attempted to experimentally estimate the internal and external reflections for an opal glass diffuser. Additional investigations by

Corresponding author address: Stanford B. Hooker, Laboratory for Hydrospheric Processes, NASA Goddard Space Flight Center, Greenbelt, MD 20771.
E-mail: stan@ardbeg.gsfc.nasa.gov

Berger (1958, 1961) refined the laboratory procedures, and Westlake (1965) gave detailed explanations for the internal and external reflection contributions. Although there were aspects of the protocols used in these early investigations that are no longer considered appropriate, many of the primary elements were properly recognized.

- A source of constant light flux is needed to vertically illuminate a diffuser at the bottom of a water vessel that has been blackened (and perhaps roughened) with waterproof *dull* black paint to minimize reflections.
- The measurements must be made in a dark room, with baffles and screens used to (a) eliminate diffuse light originating from the light source and (b) to illuminate an area only slightly larger than the diffuser.
- In-air and in-water measurements are required, and the latter should be made using optically clear water (frequently interpreted to mean tap water) or pure (distilled) water.
- A variety of water depths above the diffuser are measured, but they must exceed a so-called critical depth, $z_c = 0.9\ R_d$, where R_d is the radius of the diffuser.
- Air bubbles must be minimized, because they can create *conspicuous* bright patches, and contamination from soluble *coloring matter*, perhaps derived from the components placed in the water vessel, can influence the properties of the water being used and cannot be removed by filtering the water.

A comprehensive description of a protocol for a more modern Plexiglas diffuser was given by Smith (1969) and recommended the use of a collimated beam as a light source to avoid changes in the flux reaching the collector when the water depth changed. The study presented here is concerned with more recent diffuser designs and laboratory protocols. For the latter this means the incremental, or what is now referred to as the traditional, method. The traditional method has been in use for the past 25 yr, and originated with the protocol revisions suggested by Aas (1969) and communicated more widely by Petzold and Austin (1988). They all advocated using a lamp as a light source and including a geometric correction factor as a function of the lamp-collector distance, incremental changes in the water depth, and the water refractive index.

Mueller (1995) used the traditional method to analyze Plexiglas and Teflon diffusers for several radiometers from the same manufacturer. At any given wavelength, the immersion factors had a standard deviation (σ) between collectors that typically ranged from 3% to 5%, with total variations at some wavelengths as large as 10%. More recently, Zibordi et al. (2004) investigated the immersion factors for nine OCI-200 sensors, manufactured over a 7-yr time period by Satlantic, Inc. (Halifax, Nova Scotia, Canada), as part of the eighth SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-8). The sensors, which had identical (nominal) center wavelengths, were characterized at three different facilities (including the manufacturer's) using virtually the same traditional method. One of the radiometers was selected as a so-called reference sensor and was measured more frequently than the others.

The SIRREX-8 data showed intralaboratory repeatabilities, based on multiple characterizations of the reference radiometer, that ranged from about 0.3% to 0.6%. Interlaboratory uncertainties, evaluated with data from the nine common radiometers, showed average values lower than 0.6%. Typical $I_f(\lambda)$ values, constructed from quality-assured averages of the sensors, were less than the values supplied by the manufacturer (except one red wavelength), but were approximately within the range of variability established by Mueller (1995): more than 10% in the blue domain, and approximately 2%–6% in the green and red regions. The SIRREX-8 activity also demonstrated the inefficiencies of the traditional method: (a) sensor trial times were very long, requiring 100–330 min; and (b) the water tanks were large with water volumes measured in hundreds or thousands of liters. The lengthy experimental time limited the number of sensors characterized per day to two to five, while the large tanks required spacious work rooms, a significant capability to deal with the large amounts of water, and irreproducible volumes of water (between laboratories).

As a separate inquiry, alternatives to the traditional method (section 2) were proposed and tested with specific experiments interspersed with those designed to meet the SIRREX-8 intercomparison objectives. The new methods centered around decreasing the amount of time to execute an instrument trial and reducing the size of the experimental apparatus (specifically, the water vessel). The latter was achieved by refining the capabilities of the Compact Portable Advanced Characterization Tank (ComPACT), which had already been built for working with immersed sensors (section 3). The time efficiency was achieved primarily by making a small change to the traditional method (section 4a), and then refining the generalized protocol for the ComPACT apparatus (section 4b). The data processing requirements for the new methods share many elements with the traditional method, and the results from the use of these new capabilities (section 5) suggest they are sufficiently accurate to replace the traditional method (section 6).

2. The traditional method

The traditional method for measuring immersion factors was incorporated into the SeaWiFS Ocean Optics Protocols (hereafter referred to as the Protocols) for calibration and validation activities (Mueller and Austin 1992). The subsequent revision (Mueller and Austin 1995) and refinements to the Protocols (Mueller 2000, 2002, 2003) have not significantly changed the methodology except to note that so-called class characterizations (wherein an entire series of sensors with the same diffuser design are assigned identical immersion factors) are not suitable for calibration and validation activities; each sensor must be characterized individually. Although the latter was first noted by Mueller (1995), the protocol change did not occur until it was confirmed by the SIRREX-8 activity (Zibordi et al. 2002).

The traditional method involves a relatively simple procedure and a small number of components (Mueller and Austin 1995). A lamp of suitable wattage is needed to provide a flux of light at all sensor wavelengths well above dark values, with the appropriate baffling and apertures to minimize diffuse light contributions into the tank. A lamp with a small filament is preferred, because it better approximates a point source, and a regulated power supply should be used to ensure the lamp flux is stable over the characterization time period. The water vessel should include a removable aperture sized or adjusted to ensure the lamp flux projects onto an area that is only slightly larger than the area of the diffuser. The interior of the tank must be *flat* black and contain a sensor support system permitting an accurate horizontal leveling and vertical alignment of the sensor.

An accurate system for determining the depth of water above the diffuser, which is changed in increments (usually by draining, because turbulence is minimized), is also required. The refinements that were published in the revisions to the Protocols included (a) making it clear that the in-air measurement is only made when the diffusers are dry; (b) using a minimum water depth of 5 cm, a maximum water depth of 40–50 cm, and a water depth increment of 5 cm; and (c) repeating the measurement procedure with a different lamp-to-diffuser distance to verify an appreciable uncertainty does not affect the results.

a. Optional components

The traditional method does not include detailed specifications about several important elements, for example, the water surface (whether or not it should be kept free of floating particles) or the actual water to be used (tap water, seawater, and pure water are all possibilities). A variety of additional components (here considered *optional* for classification purposes) have proved useful by various investigators: (a) a water filter to trap particles when the tank is filled; (b) a lamp screen (1000-W lamps are frequently used and the amount of radiation is harmful); (c) intermediate apertures to ensure the light flux reaching the diffuser is as direct a beam as possible; (d) a sensor to detect anomalies in the emitted flux, which can be achieved with another radiometer, or with a digital voltmeter (DVM) measuring the voltage across a precision lamp shunt; (e) a fan to cool the lamp screen and prevent heat buildup on the monitoring sensor (if used); (f) a pump to decrease the time needed to empty the tank (particularly useful for repetitive trials); (g) a wet–dry vacuum for keeping the water surface as clean as possible; (h) an inspection port in the tank lid permitting a visual inspection of the tank interior (this can also be used for maintaining the quality of water surface with the wet–dry vacuum); and (i) a structure to isolate the sensor from the bottom of the tank (i.e., increase the height of the sensor above the turbulence around the fill and drain ports as well as any reflections from the tank bottom). Figure 1 presents all of the equipment discussed with the traditional method.

b. Traditional data processing

When an irradiance sensor is illuminated, the so-called raw optical data at each wavelength are recorded as digitized voltages $V(\lambda)$ in counts. Each sample is recorded at a specific time t_i, which also sets the water depth z. Raw irradiance data are typically converted to physical units using a calibration equation of the following form:

$$E_{\mathrm{cal}}(\lambda, t_i) = C_c(\lambda) I_f(\lambda) E(\lambda, t_i), \tag{1}$$

where $E_{\mathrm{cal}}(\lambda, t_i)$ is the calibrated irradiance, $C_c(\lambda)$ is the calibration coefficient (determined during radiometric calibration of the sensor), and $E(\lambda, t_i)$ is the net signal detected by the radiometer while illuminated. In most cases,

$$E(\lambda, t_i) = V(\lambda, t_i) - \overline{D}(\lambda), \tag{2}$$

where $\overline{D}(\lambda)$ is the average bias or dark voltage measured during a special dark measurement with the caps on the radiometer. In some cases, dark voltages are replaced by so-called background or ambient measurements, so any illumination biases can be removed along with the dark correction (Hooker et al. 2002). For the purposes of this study, references to the dark measurement represent the appropriate choice of the three possibilities.

Deriving the calculation for the immersion factor be-

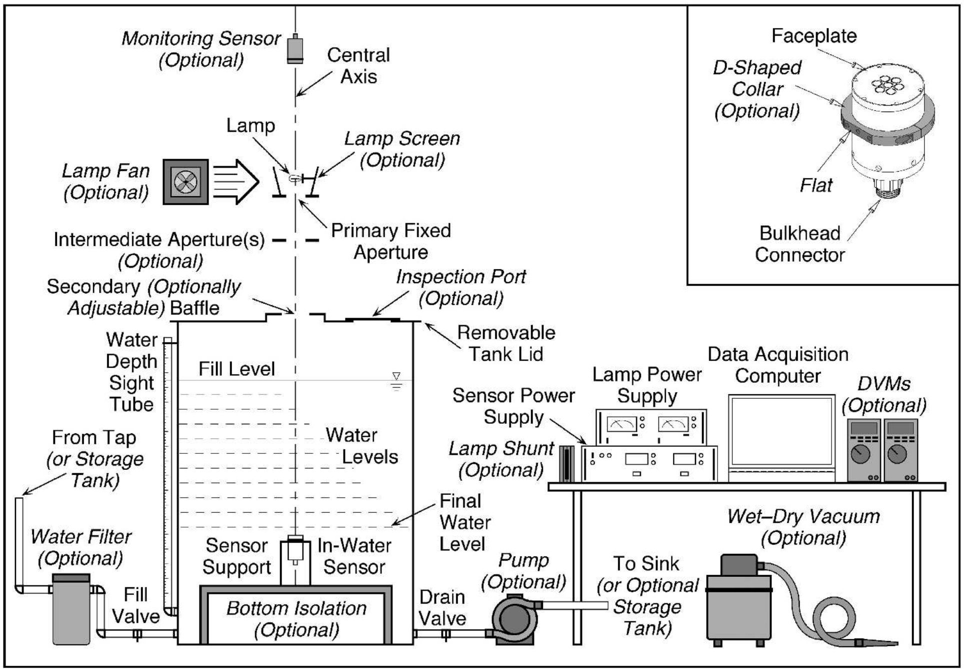

FIG. 1. The traditional laboratory setup used for characterizing the immersion factor. Optional equipment is shown in gray and labeled with a slanted typeface. In some cases, the water from the tank is drained into a sink, while for others it is pumped into a second tank (not shown) and reused in subsequent experimental trials. Alignment of the components along the central axis is best accomplished using a laser, but visual techniques (using projected shadows) have proved reliable. For most experimental systems, the water depth above the diffuser is determined using a *sight* tube mounted on the exterior of the tank. The tube is usually a clear plastic pipe with a long (adhesive) metric ruler attached to it. The water level that just begins to cover the diffusers is noted as the *null depth point*, and all subsequent readings are differenced with respect to the first reading to yield the depth of water above the diffusers. One of the most important time-saving devices is to use a pump to drain the tank, and if the water is stored in another tank between trials, to also fill it. The total time savings for the latter might be minimal, however, because some time is needed for the water to become quiescent and for bubbles to dissipate. It is appealing to maintain the angular orientation of the sensor during repetitive trials, and this is easily accomplished using the position of the flat part of a so-called D-shaped collar (the inset shows a Satlantic OCI-200 sensor fitted with a D-shaped collar). Note that the OCI-200 uses an array of diffusers, so references in this study to "a diffuser" should be interpreted as a "diffuser array" for the OCI-200 and similar designs. During SIRREX-8, the monitoring sensor was required to ensure that the only data used corresponded to a constant light flux from the lamp. In addition, two facilities used a shunt resistor in series with the lamp, as an additional means to monitor the light source stability.

gins with a simple requirement: the calibrated irradiance measured in air (indicated by $z = 0^+$), and transmitted through the air–water interface, must equal the value measured at null depth (denoted by $z = 0^-$). Using (1) with this requirement, but replacing time with depth, gives

$$E_{\mathrm{cal}}(0^-, \lambda) = E_{\mathrm{cal}}(0^+, \lambda) T_s(\lambda), \tag{3}$$

where $T_s(\lambda)$ is the transmittance of the air–water interface to downward irradiance. Substituting (1) into (3), while remembering $I_f(\lambda) = 1$ for an in-air measurement, and removing common terms (the calibration coefficient), yields

$$I_f(\lambda) = \frac{E(0^+, \lambda)}{E(0^-, \lambda)} T_s(\lambda). \tag{4}$$

The formulation in (4) assumes the $E(0^-, \lambda)$ values are exact, so if a small lamp (approximating a point source) is used, a geometric correction factor $G(z, \lambda)$ is applied to account for the change in solid angle as a function of the water depth and the distance between the lamp and the diffuser (Aas 1969; Petzold and Austin 1988):

$$G(z, \lambda) = \left[1 - \frac{z}{d}\left(1 - \frac{1}{n_w(\lambda)}\right)\right]^{-2}, \tag{5}$$

where d is the distance of the lamp source from the diffuser surface and $n_w(\lambda)$ is the index of refraction of the water, which depends on the salinity and temperature of the water (omitted here for brevity). The $G(z, \lambda)$ terms are used in the least squares fit of the logarithms of the in-water irradiance data; that is, the linear regression is computed using the $\ln[E(z, \lambda)/G(z, \lambda)]$ values. The only remaining unknown is the transmittance of the air–water interface to downward irradiance, which is computed using the Fresnel reflectance equation, $T_s(\lambda) = 4n_w(\lambda)\,[1 + n_w(\lambda)]^{-2}$.

3. The ComPACT apparatus

The ComPACT apparatus was conceived to perform tests on immersed radiometers, either in the field (measuring the response of biofouled sensors while wet) or in the laboratory (characterizing immersion factors). The original concept was to mate a small (portable) water vessel with the SeaWiFS Quality Monitor (SQM), which can be used in the laboratory or field (Johnson et al. 1998), as a portable and very stable light source (Hooker and Aiken 1998). The final ComPACT design has two light source options: (a) a lamp, which requires an adjustable aperture, or (b) the SQM, which requires a custom-built adapter.

Complete mechanical details of the ComPACT apparatus are presented in van der Linde (2003), so only a brief summary is given here. The cylindrical water vessel is approximately 45 cm long with a 10.2-cm internal diameter. The bottom is shaped to accommodate an OCI-200 sensor, which is kinematically mounted with small wing nuts. The wing nuts tighten small plates, which overhang a D-shaped clamp (Fig. 1), and affix the radiometer against the bottom. In addition to permitting an easy installation of the sensor, the D-shaped clamp also allows an accurate repositioning of the sensor during successive trials. An o-ring at the radiometer–tank interface prevents leaks. The interior of the tank is lined with 23 (equidistant) knife-edge baffles. A series of tapped holes, spaced 5 cm apart, provide an accurate control of the water level. Stainless steel cap screws, with o-rings under the caps, are used to open and close the holes. All components of the ComPACT apparatus are anodized dull black to minimize reflections.

4. Advanced methods

Although not usually stated in protocol descriptions, all methods require the collection of the appropriate number of samples at constant flux levels, which respect the signal-to-noise ratio of the instrument being characterized. For the sensors used here, all channels recorded at least 100 digital counts above dark values and more than 500 data records were acquired. Illumination constancy was achieved with a power supply stability to within 0.1%.

a. The continuous method

An alternative for characterizing immersion factors takes advantage of having a pump (with an almost constant discharge rate) to continuously drain the water vessel. For a cylindrical tank, the water depth can be approximated as a linear function of time as the water is pumped out. The total execution time is about 40 min for a 350-L tank, which is considerably shorter than the 100 min needed for the traditional method. The so-called continuous method was implemented as a modification to the traditional method, so substantial aspects of the two methods are identical. Once the sensor is properly aligned and leveled in the tank, in-air data are recorded for 3 min, and the DVM monitoring voltages are logged. The tank is filled until water begins to wash over the diffusers, and the reading on the water depth meter is noted as the null depth point. The tank is filled until the water depth above the diffusers is 50 cm; while the tank is being filled, any air bubbles forming on or near the diffusers are removed. The water surface is skimmed repeatedly with a wet–dry vacuum, and the DVM monitoring voltages are logged.

The distinctive aspects of the continuous method are associated with draining the tank. The pump is turned on, the time recorded, and sensor data are collected continuously as the tank empties. When the water level reaches the null depth point, the time is recorded, data acquisition is halted, and the DVM voltages are logged. The pump is used to lower the water depth below the sensor, the diffusers are dried using clean compressed air and lint-free tissue, and a second set of in-air data from the sensors are recorded. The collection of dark data completes the protocol.

b. The ComPACT method

The basic elements of the ComPACT protocol also involve many elements associated with the traditional method, for example, properly leveling, aligning, and adjusting the components. These and many other finer details are not recounted here, because they are provided by Zibordi et al. (2003). A summary of the successive steps involved is as follows:

1) collect dark data for both the monitoring and in-water sensors, and record the DVM monitoring voltages;

2) collect the in-air measurement with the diffusers completely dry;
3) fill the tank with pure water (produced from a process using bidistillation of the source water plus multiple purification stages, e.g., so-called Milli-Q water), and then overfill it to remove any floating particles;
4) remove any air bubbles, which form at or near the sensor diffusers, as well as on the edges of the baffles;
5) decrease the water depth in the tank using the drain hole below the current water level;
6) collect data from the in-water and monitoring sensors, and record the DVM monitoring voltages;
7) repeat steps 5 and 6 until data are collected with the lowest water depth; and
8) remove the sensor and dry the interior tank surfaces with clean compressed air.

A complete ComPACT measurement sequence typically lasts about 40 min.

The relevant parameters and quantities defining the $I_f(\lambda)$ determination through the ComPACT protocol are given in Table 1, which are compared with the equivalent values for the traditional methods used in SIRREX-8. The most significant aspects of the ComPACT method are (a) the small size of the tank and, thus, the small volume of water involved (3 L versus more than 3500 L), (b) the rapid execution time (less than 1 h rather than as much as 2–6 h), and (c) the use of a reproducible type of water (pure water rather than local tap water or aged seawater).

5. Results

The primary objective of this study is to evaluate alternative—more time efficient and more reproducible—methods for characterizing the immersion factor. Although it is straightforward to establish time efficiency, the processes influencing uncertainties require separate experiments wherein a range of parameters are varied and compared. The comparisons are formulated here by computing a relative percent difference (RPD), ψ, defined as

$$\psi = 100\,\frac{Y - X}{X}, \tag{6}$$

where X is a reference measurement and Y is the measurement under evaluation.

The validation of the alternative methods is accomplished by comparing them to the traditional method.

TABLE 1. A comparison of the principal methodological parameters used by the different SIRREX-8 laboratories executing the traditional method (Zibordi et al. 2002) vs the ComPACT method. The number in parentheses beside the tank radius is the tank radius divided by the radius of the area enclosing the OCI-200 diffuser array, and the value given next to the depth increment is the number of depth intervals measured.

Parameter(s) and units	Traditional methods			ComPACT
Tank volume (L)	780	350	3, 527	3
Tank radius (cm) (radii ratio)	40.5 (12.7)	40.0 (12.6)	45.0[a] (14.2)	5.1 (1.6)
Interior tank obstructions	None	Small	Large	Sidewall baffles
Water type	Tap water	Tap water[b]	Seawater	Pure water[c]
Water age (days)	<1	~2–3	>300[d]	<1
Surface cleaning	Vacuum	Vacuum	Soap[e]	Overfill
Lamp power (W)	400	1000	1000	1000[f]
Lamp filament	Small	Large[g]	Large	Large[g]
Lamp-to-diffuser distance	86.0[h]	105.0	100.0	125
Depth increment (cm; No.)	2.5 (15)	2.5 (13)	5.0 (10)	5.0 (7)
Characterization time (min)	120	100	330[i]	40

[a] The smaller dimension of the 90 cm × 123 cm tank.
[b] Demineralized (by filtration) tap water, with a typical resistance of 5–8 MΩ.
[c] Milli-Q water.
[d] Aged for about 1 yr, occasionally chlorinated, intermittently contaminated with small quantities of soap, and coarsely filtered each time the tank is drained.
[e] A small amount of soap is added to the water, and the soap molecules spread on the water surface to a monomolecular layer, thereby sweeping away any surface particles (large particles are removed with a dip net). It is assumed that the soap layer does not modify the transmittance of the water surface.
[f] A 4.0- and 4.2-A lamp current was used for upward (E_u) and downward (E_d) irradiance sensors, respectively.
[g] The filament is U shaped, which allows the outer glass envelope to be relatively small.
[h] A second distance of 100.0 cm is an option.
[i] Includes 300 min for tank filling and settling.

Specifically, the $I_f(\lambda)$ values from the continuous or ComPACT method and a specific sensor (Y) are differenced with respect to the $I_f(\lambda)$ values for that sensor from the traditional method (X). Average RPD values are useful in evaluating overall performance characteristics of the methods, and are produced by averaging all the individual $I_f(\lambda)$ from the set of sensors being measured or by averaging across all wavelengths (which were identical for all the sensors).

A summary of the sensors characterized during SIRREX-8 is presented in Table 2. The mixture of new and old sensors covered a 7-yr span of instrument production, with the reference sensor (Eu130) being one of the newest. Although the field sensors were subjected to a diverse set of field campaigns and shipping circumstances, they were all cared for diligently. This included the use of double-packed (box within a box) professionally designed shipping containers, regular calibrations (either at a recurring frequency or before and after field deployments), custom-made instrument stands to minimize the likelihood of instrument damage in the field, etc. A visual inspection of the diffusers for the field instruments showed they were all in very good condition.

TABLE 2. The radiometers used during SIRREX-8. The sensor codes are formed from the first two letters of the measurement type, plus a three-digit serial number. The manufacturing dates are based on the first time the instruments were calibrated by the manufacturer in the configuration they were used for SIRREX-8. The Eu047 sensor was modified to have low saturation levels, with respect to the standard in-water configuration, so it could be used with sources emitting low light levels. It was also only used in laboratory experiments, so it did not have any of the inevitable diffuser degradation associated with field instruments. To further determine whether or not field use resulted in any unusual aging properties, two new sensors, Ed161 and Eu162, were also measured. All the sensors had identical (nominal) center wavelengths: 412, 443, 490, 510, 555, 665, and 683 nm.

Sensor code	Sensor type	Date of manufacture	Instrument notes
Ed015	$E_d(\lambda)$	Sep 1994	Oldest sensor
Ed040	$E_d(\lambda)$	Mar 1996	
Eu047*	$E_u(\lambda)$	Jun 1996	Special sensor with the greatest sensitivity
Eu048	$E_u(\lambda)$	Jun 1996	
Ed050	$E_d(\lambda)$	Jun 1996	
Ed071	$E_d(\lambda)$	Apr 1997	
Ed097	$E_d(\lambda)$	Jun 1998	
Eu098	$E_u(\lambda)$	May 1998	
Eu109	$E_u(\lambda)$	Jul 1998	
Eu130	$E_u(\lambda)$	Jul 1999	*Reference* sensor (measured most frequently)
Ed161*	$E_d(\lambda)$	Sep 2001	Manufactured right before SIRREX-8
Eu162*	$E_u(\lambda)$	Sep 2001	Manufactured right before SIRREX-8

* Only used in the laboratory and never in the field.

a. Validation of the alternative methods

Experiments with the continuous methods were usually made immediately before or after a traditional method trial, so temporal differences in many of the parameters would be minimized. Figure 2 presents the comparison of all the sensors measured using the two methods for which data collection occurred during the same day. The mix of sensors is not the complete set presented in Table 2, because of the focused objectives of this study, but a significant majority are included. The histogram of RPD values between the two methods (inset in Fig. 2) is very nearly a Gaussian distribution, with an insignificant positive bias (the average RPD is 0.1%). Note the anomalous behavior of the Ed015 sensor and that the outer edge of the range of variance is defined primarily by the Ed040 and Ed050 sensors.

The comparison of the ComPACT and traditional methods is given in Fig. 3, and is also based on using experimental trials executed in as small a time difference as possible. The mix of sensors is different than those used with evaluating the continuous method, but five are common to both, and between the two types of experiments, all 12 sensors (Table 2) are represented. The histogram of RPD values (inset in Fig. 3) has a significant central peak, but a distorted Gaussian distribution—there is a small net positive bias; the average RPD is 0.2%, which is only a little larger than the bias seen with the continuous method.

Regardless of the characterization methodology, Figs. 2 and 3 show sensor-to-sensor differences can be substantial and are not exclusively associated with older instruments. The Ed040 and Ed050 sensors define much of the outer range in variance, but the newer Eu130 and Ed161 sensors have channels that perform similarly. A generalized description of the spectral variance is that many sensors exhibit both lesser and greater variability, rather than consistently minimal or maximal variability. The good agreement between the alternative and traditional methods is also seen in the repeatability of individual sensors. The average repeatability (across all wavelengths) of the traditional, continuous, and ComPACT methods is 0.5%, 0.3%, and 0.5%, respectively (defined by 2σ in the immersion factors for repetitive characterizations of the same sensor). Although not all the sensors were measured more than once, the majority of the sensors were, and the reference sensor (Eu130) was characterized several times for all three methods.

Another way of intercomparing the methods is to

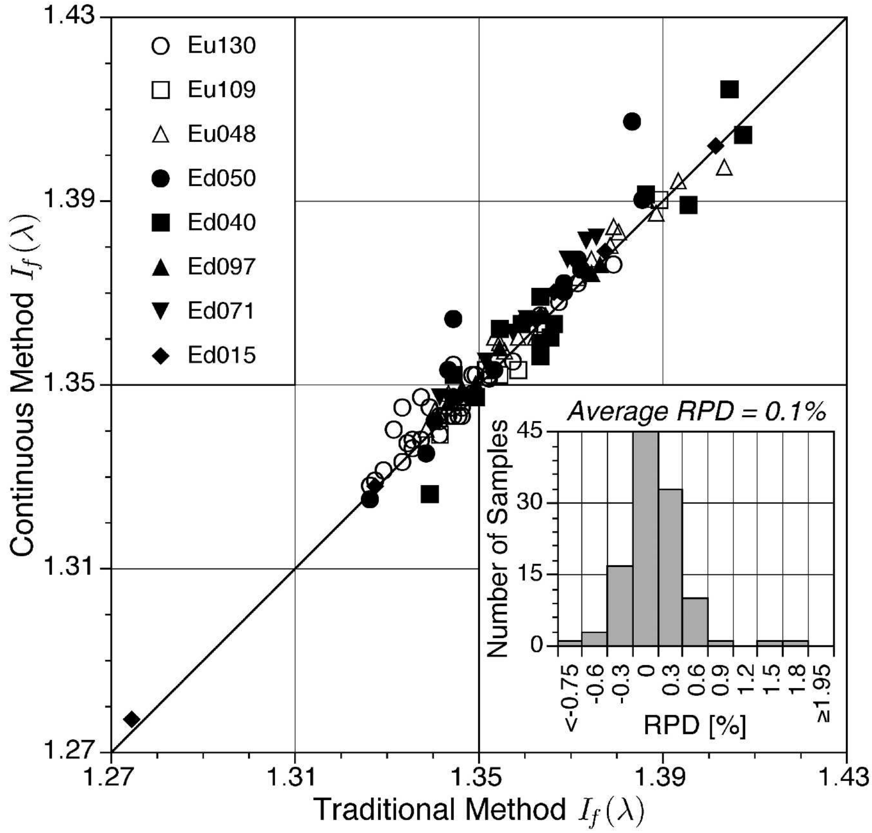

FIG. 2. A comparison of the traditional and continuous methods for characterizing $I_f(\lambda)$. The inset shows the histogram of RPD values between the methods, with the traditional method used as the reference in the RPD calculations.

compare the immersion factors from all sensors on all days—not just the ones measured on the same day—to see how similar the resulting class characterizations are with the three different methods. Although class characterizations are no longer recommended for calibration and validation activities, they are still very common, because of the expense involved with determining the response of individual sensors. Table 3 presents a comparison of the class characterizations using the traditional, continuous, and ComPACT methods. The similarity of the results does not depend on including or excluding the (apparently) anomalous Ed015 sensor. The RPD analysis shows the continuous and ComPACT methods almost always intercompare with the traditional method to within less than 1% (with one exception), with some indication of biases: the continuous trusted results are all slightly negative, and the ComPACT results are mostly positive.

Considering the maximum uncertainties in $I_f(\lambda)$ values as defined by 2σ in the spectral estimates across all the OCI-200 sensors, the traditional, continuous, and ComPACT methods give very similar results, which vary between 1.4% and 3.4%, 0.8% and 3.1%, and 0.7% and 3.3%, respectively. When combined with the previous nearly unbiased levels of agreement between the methods on a sensor-by-sensor basis (which ranged from 0.1% to 0.2%, on average), plus the excellent repeatability of each method (approximately 0.5%), the majority of the uncertainty comes from the differences between the instruments. Furthermore, the traditional, continuous, and ComPACT methods are indistinguishable from one another, and all three methods are equally capable to within the uncertainties in the instruments and methods.

b. *Sources of uncertainties*

The method intercomparison results were based on a concerted effort to implement the three methods while minimizing any sources of perturbations or uncertainties. Such an approach does not permit a thorough description of the degree of degradation that can be expected from each source of uncertainty. Consequently, specific experiments were conducted, so the magni-

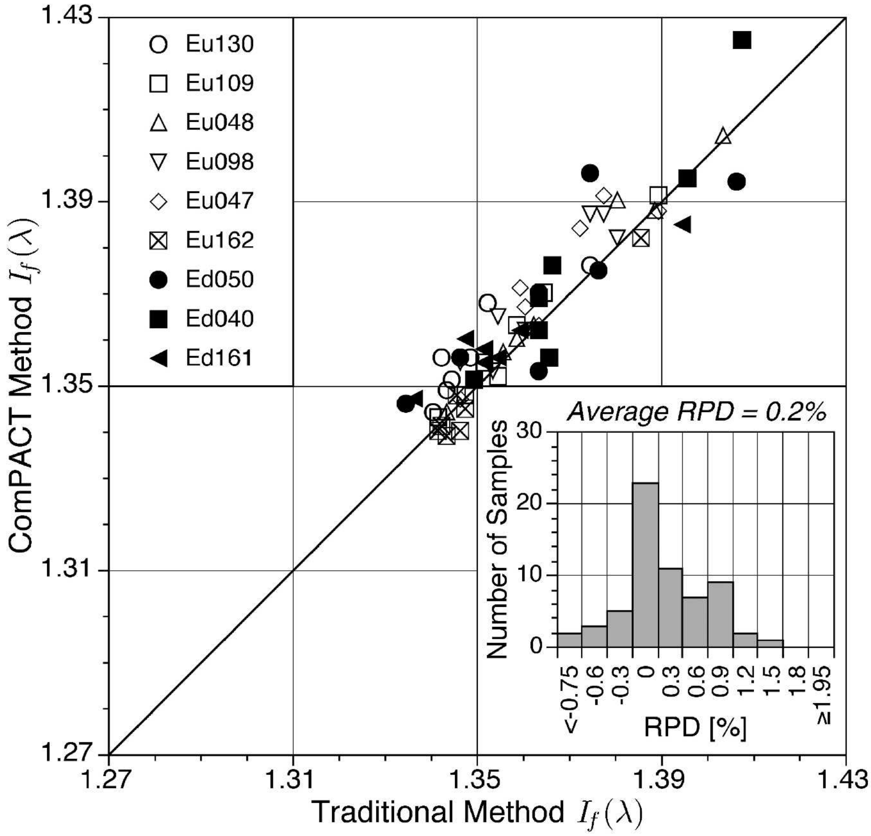

FIG. 3. A comparison of the traditional and ComPACT methods for characterizing $I_f(\lambda)$. The inset shows the histogram of RPD values between the methods, with the traditional method used as the reference in the RPD calculations.

tudes and biases of the uncertainties could be understood.

1) SURFACE PARTICLES

One of the most visually discernible uncertainties is the quality of the water surface, because the unidirectional illumination of the water vessel makes it easy to see floating particles on the water surface or in the air (a persistent source of contamination). Surface particles reflect light and lower the flux reaching the submerged sensor, which artificially increases the immersion factor (4). Figure 4 presents a set of experiments

TABLE 3. A comparison of the spectral overall averages—or class characterizations—of the immersion factors for the SIRREX-8 sensors as measured with the traditional, continuous, and ComPACT methods. The RPD values are computed using the former as the reference with respect to the latter two. Following the presentation of the SIRREX-8 results (Zibordi et al. 2004), the data are split into two categories: (a) all the sensors and (b) the trusted sensors (all the sensors with Ed015 omitted). The $I_f(\lambda)$ values for the traditional method are taken directly from the full SIRREX-8 results.

	Immersion factor						RPD (%)			
	Traditional		Continuous		ComPACT		Continuous		ComPACT	
Wavelength (nm)	All	Trusted	All	Trusted	All	Trusted	All	Trusted	All	Trusted
412	1.349	1.355	1.353	1.349	1.360	1.355	0.3	−0.4	0.8	0.0
443	1.381	1.385	1.380	1.378	1.386	1.384	0.0	−0.5	0.3	−0.1
490	1.354	1.358	1.360	1.353	1.363	1.357	0.4	−0.4	0.7	−0.1
510	1.350	1.350	1.343	1.343	1.349	1.348	−0.5	−0.5	−0.1	−0.1
555	1.363	1.367	1.358	1.357	1.369	1.367	−0.3	−0.8	0.4	0.0
665	1.355	1.370	1.355	1.364	1.364	1.371	0.0	−0.4	0.7	0.1
683	1.367	1.379	1.373	1.373	1.383	1.384	0.4	−0.4	1.2	0.4

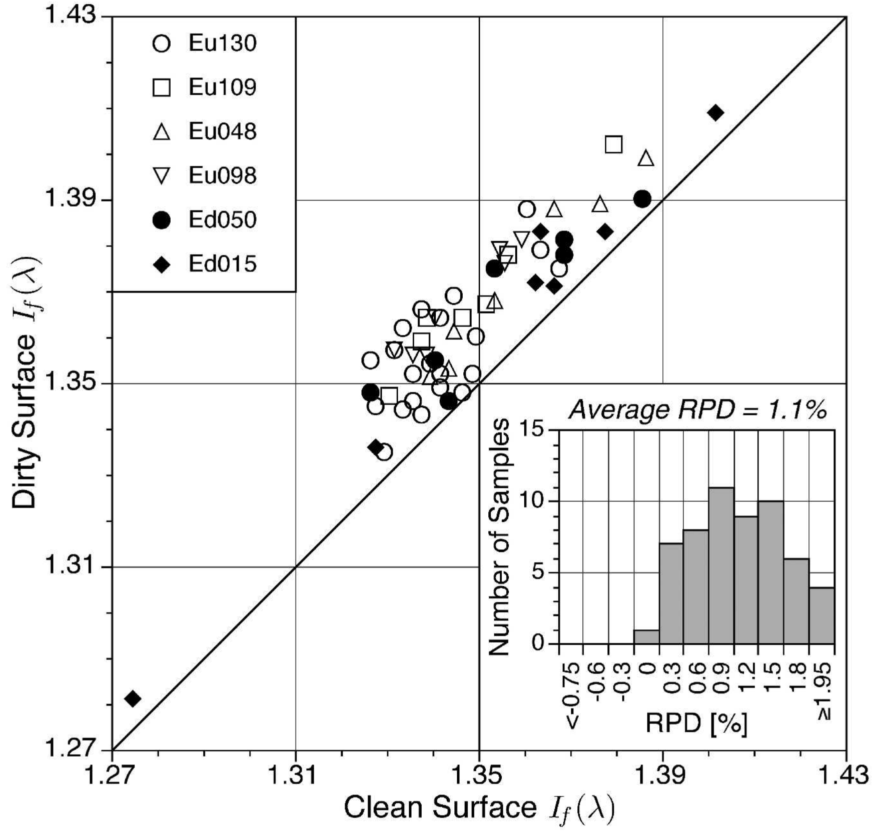

FIG. 4. A comparison of $I_f(\lambda)$ values determined with the traditional method for clean and dirty surface conditions. The former was accomplished using a wet–dry vacuum, whereas the latter was established by ignoring the cleaning requirement. The inset shows the histogram of RPD values between the two types of measurements, with the clean surface data used as the references in the RPD calculations. Multiple trials were involved with the Eu130 sensor, so these data appear more often.

wherein instruments were characterized with the traditional method when the water surface was both dirty and clean. The latter was accomplished primarily by using a wet–dry vacuum to repeatedly skim the water surface immediately prior to data collection. For these data, the clean surface results are used as the reference measurement in the RPD calculation (6). There is a significant increase in the $I_f(\lambda)$ values associated with a dirty surface, and the average bias can be greater than 1%.

The results from the experiments evaluating different levels of surface contamination are shown in Fig. 5. The data demonstrate four important aspects of this problem: (a) the uncertainty can be on the order of a few percent, (b) any amount of surface cleaning reduces the uncertainty, (c) surface vacuuming can reduce the uncertainty to within the measurement repeatability (approximately 0.5%), and (d) the spectral properties are essentially *white* (i.e., they are mostly independent of the wavelengths considered here).

2) In-water absorption and scattering

The water surface is an easily inspected part of the filled water vessel. Tanks equipped with inspection ports also permit the column of water in between the surface and the sensor to be viewed, and it is not unusual to see debris floating around in the tank. In many cases, the origin of the debris is the air above the tank, but in other cases it is from wherever the water is taken. Water filters can help remove particles, but there is a diminishing rate of return with this approach: trapping a wide cross section of particles requires a very fine filter, which reduces the flow rate and increases the length of time to conduct an experiment. Submerged particles act as scatterers and increase the amount of light reaching the sensor, which artificially lowers the immersion factor (4). The other common sources of scattered light are bubbles and reflections off the sidewall of the water vessel. Bubbles can adhere to any surface inside the tank, but they are most damaging if

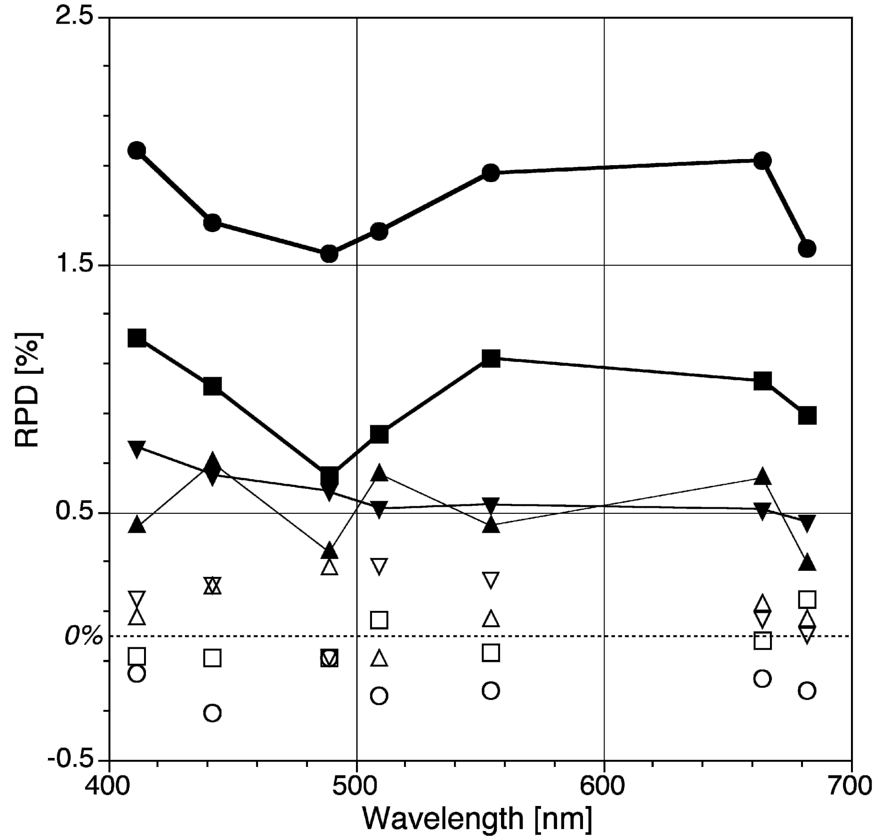

FIG. 5. The effects of progressively more aggressive surface cleaning on the uncertainty (RPD) in characterizing $I_f(\lambda)$ for the Eu130 sensor with the traditional method. The cleanest surface conditions (from extensive wet–dry vacuuming) are given by the open symbols, and the effect of cleaning a dirty surface is given by the sequence of solid symbols. The latter starts with no surface cleaning for an unrealistic water surface covered by a high concentration of particles (solid circles and solid squares) and concludes with extensive manual removal and light wet–dry vacuuming (solid triangles). The RPD values are computed by using the average of the clean surface data as the reference in (6).

they adhere to the diffusers. Although they can appear spontaneously as the water changes temperature, they are usually caused by using a pump to fill the water vessel.

The other in-water property that needs to be considered is absorption, specifically dissolved colored material that might be a part of the source water or the result of contamination from equipment placed in the tank. If the water is being stored, it is possible for absorbers and scatterers to build up over time and produce time-varying perturbations. Although naturally aging a volume of water might provide a range of water types, this is not a particularly efficient or reproducible approach. For experimental trials where the water needs to be contaminated, the ComPACT method offers substantial advantages, because small volumes can be easily prepared and the apparatus can be cleaned completely and quickly after the experiment is completed. An exploration of absorption and scattering effects was made with the ComPACT method. The addition of a large amount of a scattering agent, in this case Formazine, resulted in decreased immersion factors at all wavelengths (-5.2% on average), with maximal changes in the blue domain (approaching -10%). Conversely, the addition of an absorption agent had a small effect (-0.2% on average), even though the large amount of added material caused a noticeable change in water color.

3) DATA PROCESSING

Quantification of the uncertainties induced by the processing scheme was addressed by comparing the $I_f(\lambda)$ values determined with two different processors for the continuous method. The primary difference between the processors is one bins the data vertically to essentially convert the continuous method into the traditional method, whereas the other deals with the data as a continuum (D'Alimonte and Brown 2003). The average RPD between the processors is approximately 0.2%. Both processors make use of the same relationships for determining $n_w(\lambda)$, $G(z, \lambda)$, and $T_s(\lambda)$, so the differences are mostly explained by the implementation of a noise reduction filter in one processor, and by the binning algorithm in the other. Because the differences between the two processors are within the repeatability of $I_f(\lambda)$ determinations (about 0.5%), no additional effort was made in producing a better convergence of the results.

6. Discussion and conclusions

The evaluation of the continuous method demonstrates its full applicability in the determination of immersion factors with a significant time savings. Similarly, the evaluation of the ComPACT method demonstrates the possibility of significantly reducing the size of the tank (along with decreasing the execution time) and permitting a completely reproducible methodology (pure water). There are additional benefits with the ComPACT method, however, and these include the following: (a) the water surface can be cleaned by overfilling the tank, so the contaminated surface is quickly washed away; (b) a modest amount of space and significantly simpler waste water capabilities are needed; and (c) specialized experiments with contaminated water can be accommodated, because of the feasibility of producing small volumes of homogeneous and reproducible solutions (e.g., the addition of absorption or scattering materials, or other substances associated with a unique measurement environment).

Within the discussions of uncertainties and the ultimate selection of an immersion factor method, it is important to remember the analytical approach adopted here can be considered conservative, because the uncertainties are not shared equally—all the differences in the new versus traditional comparisons were ascribed to the new methods. This is a consequence of there being no absolute truth associated with the entire process, so the traditional method was selected as the reference for evaluation purposes. Consequently, if a new method satisfies the general protocol, and produces results within the uncertainties of the accepted method, there is no reason to ignore it, particularly if it provides demonstrable advantages (like the ones listed above).

The concept that there is no absolute truth in characterizing the immersion factor is an important one. There is nothing that can be purchased from a standards laboratory that will allow an investigator to compare the experimental results with a set of known values. The *answer* is achieved experimentally by following an accepted protocol as accurately as possible. The summary of the validation results produced for the different methods showed a significant convergence of the continuous and ComPACT methods with the traditional method (Figs. 2 and 3 plus Table 3). The maximum uncertainties in the immersion factors for the OCI-200 series of radiometers determined with the three methods and defined by 2σ in the spectral estimates across all the sensors were all approximately 1%–3%.

The average repeatabilities for single-sensor characterizations (across the seven wavelengths) of the three methods were very similar and approximately 0.5%. This means the majority of the uncertainty is caused primarily by sensor-to-sensor differences in the immersion factor, and not by differences between the methods. The achieved level of convergence and repeatability indicates the individual sources of uncertainty were properly minimized (there was no notable bias in the results like was seen with the surface reflectance experiments in Figs. 4 and 5). The absence of significant biases does not mean subtle differences were not discerned, however. The $I_f(\lambda)$ values determined with the ComPACT method, for example, are a little higher than those obtained with the other methods. This result is most likely explained by the use of pure (Milli-Q) water with the ComPACT method, which is cleaner than the demineralized tap water used with the other methods and, thus, yields more accurate and slightly higher $I_f(\lambda)$ values.

Irrespective of the chosen immersion factor method, the Protocols do not explicitly address the importance of all the sources of uncertainty encountered during a sensor characterization activity. Indeed, a continuing paradox in the Protocols is that the laboratory method is supposed to be executed with tap water, but the sensor is almost always used in seawater. Although one SIRREX-8 facility used seawater, the water was sufficiently irreproducible (Table 1) that a controlled experiment to investigate the effects of seawater versus pure water could not be made. This was easily accomplished with the ComPACT apparatus (Zibordi et al. 2003), however, as were a number of other investigations concerning isolated aspects of the experimental process.

A summary of the investigation of individual sources of uncertainties in characterizing immersion factors is presented in Table 4. Some of the uncertainties are on the order of the level of repeatability (approximately 0.5%) and rather uniformly increase or decrease the immersion factor across all wavelengths (positive and negative RPD values, respectively). The exceptions are

TABLE 4. Typical uncertainties for characterizing immersion factors for a variety of experimental conditions. All the experiments were conducted with the ComPACT method except the first one. The ComPACT uncertainties are estimated using RPD values computed by comparing the results for each experiment associated with the source of uncertainty under investigation with the results obtained for pure water with the uncertainty source absent (the latter is the reference in the RPD calculations). The overall bias for each uncertainty is given by the spectral average.

	Wavelength (nm)							
Source of uncertainty	412	443	490	510	555	665	683	Spectral avg
Dirty surface[a]	1.1	1.0	0.8	0.9	1.0	1.0	0.8	0.9
Saltwater[b]	0.4	0.5	0.6	0.3	0.4	0.4	0.6	0.5
Sidewall reflections[c]	0.1	0.2	0.2	0.1	−0.1	0.0	0.0	0.1
High absorption[d]	−0.2	−0.2	0.2	−0.2	−0.4	−0.4	−0.1	−0.2
Tap water[e]	−0.5	−0.4	−0.1	−0.7	−0.2	−0.2	−1.0	−0.4
Aged pure water[f]	−0.5	−0.6	−0.6	−0.5	−0.6	−0.5	−0.4	−0.5
Aged tap water[g]	−1.3	−0.9	−0.9	−1.0	−1.0	−1.1	−1.0	−1.0
Bubbles[h]	−1.2	−0.9	−2.2	−1.1	−2.4	−0.9	−1.6	−1.5
High scattering[i]	−9.9	−6.7	−5.0	−4.7	−4.3	−3.1	−3.0	−5.2

[a] The average of the dirty surface conditions (solid symbols) in Fig. 5.

[b] Produced by filtering (0.22-μm pore size) real and synthetic seawater (Zibordi et al. 2003): (a) seawater from the northern Adriatic Sea and (b) 3.5% pure sea salt and aquarium sea salt added to volumes of pure water.

[c] Created by removing the adjustable aperture above the water vessel to ensure the baffled sidewalls were illuminated.

[d] Produced by adding colored dissolved matter to pure water. Absorbance at 400 nm was 0.020 (dimensionless) over a 10-cm cuvette.

[e] Demineralized (by filtration) tap water, with a typical resistance of 5–8 MΩ.

[f] Aged for approximately 3 days.

[g] Demineralized tap water aged for approximately 5 days.

[h] Produced by adding carbonated mineral water to pure water.

[i] Produced by mixing Formazine into pure water. Absorbance at 400 nm was 0.084 (dimensionless) over a 10-cm cuvette.

the uncertainties associated with sidewall reflections, tap water, and high water absorption, which have a smaller impact on the uncertainty budget (on average). Note that the former is only applicable to the ComPACT water vessel and demonstrates the baffles were effective; an alternative tank without baffling might be more susceptible to sidewall reflections.

Water quality is confirmed as an important source of uncertainty in Table 4, and the signs of the uncertainties (all negative) suggest the primary problem is an increase in suspended scatterers. The most extreme examples of scattering (bubbles and water contaminated with Formazine) show more spectral structure than the other sources of uncertainty and establish another reason for avoiding these problems. The results indicate the use of tap water or recycled (aged) water should be avoided, and the most effective method will rely on pure water manufactured immediately before use. The saltwater results suggest a pure water characterization can be rather effectively used in seawater by adding a constant 0.5% correction factor as proposed by Zibordi et al. (2002).

Positive and negative biases suggest some methods can experience compensation—the partial or complete cancellation of competing uncertainties—and produce fortuitous agreement with other methods. For example, a method using recycled tap water with no surface cleaning might agree rather favorably with a pure water method using surface skimming. Ultimately, the criteria for choosing a method should be based on ensuring that the sources of uncertainties can be individually discriminated and minimized. Furthermore, any method allowing the methodological elements to be implemented rapidly and easily is preferred over an alternative requiring a substantial amount of extra effort. Within this context of capability and efficiency, the ComPACT apparatus (perhaps equipped with a constant-flow pump) emerges as the most effective method of the ones considered in this study.

Acknowledgments. The collection and analysis of the SIRREX-8 data were supported primarily by the Coastal Monitoring Project at the JRC; the SeaWiFS and SIMBIOS projects at the NASA Goddard Space Flight Center provided additional support whenever it was required. The processing of the continuous data as a continuum of observations used software developed by Jim Brown (University of Miami) and the SeaWiFS Field Team. The JRC processing software was coded by Davide D'Alimonte. Acknowledgments are due to Dirk van der Linde for the effort placed in the design interpretation and mechanical production of the COMPACT apparatus. The investigation of uncertainties in radiometric observations has been a high priority for James Mueller; his continuing willingness to discuss and place an emphasis on the many aspects of measurement protocols is appreciated.

REFERENCES

Aas, E., 1969: On submarine irradiance measurements. University of Copenhagen Rep. 6, Institute for Physical Oceanography, University of Copenhagen, Copenhagen, Denmark, 23 pp.

Atkins, W. R. G., and H. H. Poole, 1933: The photo-electric measurement of the penetration of light of various wave lengths into the sea and the physiological bearing of results. *Philos. Trans. Roy. Soc. London,* **222,** 129–164.

Berger, F., 1958: Über die ursache des "oberflächeneffekts" bei lichtmessungen unter wasser (translated from German). *Wetter Leben,* **10,** 164–170.

——, 1961: Über den "Taucheffekt" bei der lichtmessung über und unter wasser (translated from German). *Arch. Meteor. Wien.,* **11,** 224–240.

D'Alimonte, D., and J. W. Brown, 2003: Processing of alternative methods. *New Laboratory Methods for Characterizing the Immersion Factors of Irradiance Sensors*, NASA Tech. Memo. 2003-206892, S. B. Hooker and E. R. Firestone, Eds., Vol. 26, NASA Goddard Space Flight Center, Greenbelt, MD, 21–24.

Hooker, S. B., and W. E. Esaias, 1993: An overview of the SeaWiFS project. *Eos, Trans. Amer. Geophys. Union,* **74,** 241–246.

——, and J. Aiken, 1998: Calibration evaluation and radiometric testing of field radiometers with the SeaWiFS Quality Monitor (SQM). *J. Atmos. Oceanic Technol.,* **15,** 995–1007.

——, and C. R. McClain, 2000: The calibration and validation of SeaWiFS data. *Progress in Oceanography*, Vol. 45, Pergamon, 427–465.

——, S. McLean, J. Sherman, M. Small, G. Lazin, G. Zibordi, and J. W. Brown, 2002: The Seventh SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-7), March 1999. NASA Tech. Memo. 2002-206892, Vol. 17, NASA Goddard Space Flight Center, Greenbelt, MD, 69 pp.

Johnson, B. C., P.-S. Shaw, S. B. Hooker, and D. Lynch, 1998: Radiometric and engineering performance of the SeaWiFS Quality Monitor (SQM): A portable light source for field radiometers. *J. Atmos. Oceanic Technol.,* **15,** 1008–1022.

Mueller, J. L., 1995: Comparison of irradiance immersion coefficients for several marine environmental radiometers (MERs). *Case Studies for SeaWiFS Calibration and Validation, Part 3*, NASA Tech. Memo. 104566, S. B. Hooker, E. R. Firestone, and J. G. Acker, Eds., Vol. 27, NASA Goddard Space Flight Center, Greenbelt, MD, 3–15.

——, 2000: Overview of measurement and data analysis protocols. *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 2*, NASA Tech. Memo. 2000-209966, NASA Goddard Space Flight Center, Greenbelt, MD, 87–97.

——, 2002: Overview of measurement and data analysis protocols. *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 3*, *Volume 1*, NASA Tech. Memo. 2002-210004/Rev3-Vol1, NASA Goddard Space Flight Center, Greenbelt, MD, 123–137.

——, 2003: Overview of measurement and data analysis methods. *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4*, *Volume III: Radiometric Measurements and Data Analysis Protocols*, NASA Tech. Memo. 2003-211621/Rev4-Vol.III, NASA Goddard Space Flight Center, Greenbelt, MD, 1–20.

——, and R. W. Austin, 1992: Ocean optics protocols for SeaWiFS validation. NASA Tech. Memo. 104566, Vol. 5, NASA Goddard Space Flight Center, Greenbelt, MD, 43 pp.

——, and ——, 1995: Ocean optics protocols for SeaWiFS validation, revision 1. NASA Tech. Memo. 104566, Vol. 25, NASA Goddard Space Flight Center, Greenbelt, MD, 67 pp.

Petzold, T. J., and R. W. Austin, 1988: Characterization of MER 1032. Tech. Memo. EN-001-88t, Visibility Laboratory, Scripps Institution of Oceanography, La Jolla, CA, 56 pp.

Smith, R. C., 1969: An underwater spectral irradiance collector. *J. Mar. Res.,* **27,** 341–351.

van der Linde, D., 2003: The ComPACT mechanical design. *New Laboratory Methods for Characterizing the Immersion Factors of Irradiance Sensors*, NASA Tech. Memo. 2003-206892, S. B. Hooker and E. R. Firestone, Eds., Vol. 26, NASA Goddard Space Flight Center, Greenbelt, MD, 14–16.

Westlake, D. F., 1965: Some problems in the measurement of radiation under water: A review. *Photochem. Photobiol.,* **4,** 849–868.

Zibordi, G., D. D'Alimonte, D. van der Linde, J.-F. Berthon, S. B. Hooker, J. L. Mueller, G. Lazin, and S. McLean, 2002: The Eighth SeaWiFS Intercalibration Round-Robin Experiment (SIRREX-8), September–December 2001. NASA Tech. Memo. 2002-206892, Vol. 21, NASA Goddard Space Flight Center, Greenbelt, MD, 39 pp.

——, ——, ——, S. B. Hooker, and J. W. Brown, 2003: New laboratory methods for characterizing the immersion factors of irradiance sensors. NASA Tech. Memo. 2003-206892, Vol. 26, NASA Goddard Space Flight Center, Greenbelt, MD, 34 pp.

——, S. B. Hooker, J. Mueller, and G. Lazin, 2004: Characterization of the immersion factor for a series of in-water optical radiometers. *J. Atmos. Oceanic Technol.,* **21,** 501–514.